Strategies for GIS and Public Health

Michael F Goodchild*
National Center for Geographic Information and Analysis, University of California, Santa Barbara, CA; Department of Geography, University of California, Santa Barbara, CA

Abstract

The paper is divided into three sections. The first reviews three broad trends in information technology that will affect geographic information systems (GIS) in the coming years. The second identifies four trends that are specific to GIS, including availability of new data, trends in software, developments in education, and opportunities in new hardware. The third section recommends six potential strategies for advancing applications of GIS to public health. The paper ends by suggesting an analogy between imaging the body and imaging the health of the nation through GIS display and analysis.

Keywords: public health, spatial data infrastructure, future trends, strategies

Introduction

As the conference closed, I was struck by two things: the refreshing diversity among the conference attendees, and the sheer size of this geographic information system (GIS) application domain. Many GIS conferences still have a long way to go before their audiences resemble America, but this one seemed much closer to achieving that goal. The conference showed ample evidence of the vast range of health applications of GIS in the variety of topics among the roughly 130 papers, the posters and demonstrations, and the pre-conference workshops. It showed the power of GIS for mapping, but also for gaining new insight by displaying data in new ways and by organizing data and programs geographically. Yet it seemed to me that the conference demonstrated the potential for a community of health professionals using GIS that would be as much as two orders of magnitude greater in number than the present community, given the potential for improved health as a goal of many, many different types of GIS applications.

The following three sections address the three distinct purposes of this paper:

1. To identify trends in society and in information technology that are larger in scale than GIS, and that will inevitably affect the application of GIS to health problems.
2. To identify trends that are specific to GIS, but still larger than any one application.
3. To suggest some strategies for the community interested in GIS applications to health problems.

The three sections are followed by a brief conclusion.

External Trends

The Mapping Science Committee (MSC) of the National Research Council exists as "a

* Michael F Goodchild, Dept. of Geography, University of California, Ellison 3611, Santa Barbara, CA 93106-4060 USA; (p) 805-893-8049; (f) 805-893-3146; E-mail: good@ncgia.ucsb.edu

focus for external advice to the federal agencies on scientific and technical matters related to spatial data handling and analysis," and has been instrumental in initiating the concept of the National Spatial Data Infrastructure (1). In 1997, the MSC published a report entitled *The Future of Spatial Data and Society* (2), which reported on a workshop sponsored by the committee in 1996. The report reviews trends in society at large that are likely to impact GIS in the next 15 years, and lays out a number of alternative visions for the role of GIS in society in 2010. Below, I review three of those trends, chosen because they seem to be of particular significance to applications of GIS to health problems.

Information Technology

One of the most remarkable trends of the past 15 years has followed the prediction attributed to Gordon Moore, co-founder of Intel, that processor speed would approximately double every 18 months, and that processor costs would stay approximately constant. Two things are remarkable about the prediction: first, the precise way in which the actual performance of the industry has matched it; and second, the fact that it was proposed by one of the key figures in that industry, rather than an independent observer. Over the past 15 years and since the advent of the IBM PC, the speed of a PC's central processor has increased by a factor of roughly 1,000, and costs have stayed roughly the same. Similar improvements in price and performance have occurred in other key areas of the computing industry—memory, hard disks, and CDs.

Over the years, there have been numerous speculations on the accuracy of Moore's Law, and in mid-1997 a number of articles in the popular press announced its demise. But while most speculations have focused on possible under-performance by the industry, the mid-1997 articles predicted over-performance: that speed would begin to rise at an even greater rate due to a series of key breakthroughs in chip design and manufacture. At this point in time, there seems very little reason to be pessimistic about the future of information technology.

Looking back, it is interesting to ask where all the new cycles and bytes made possible by Moore's Law have gone; certainly, there is little evidence that every GIS application is making 1,000 times the number of numerical calculations it was making in 1984. Instead, it seems that much of the new power has gone into areas that in 1984 would have been regarded as somewhat superficial: maintaining a graphic user interface with a specific "look and feel," supporting a connection to the network, and in general making the system easier to use. It seems that every new version of the operating system is friendlier and more visual, but also demands more resources. One might even consider a corollary to Moore's Law: that such non-essential aspects of computing will consume a constant proportion of the increasing resources. Certainly, and despite continued concerns about the difficulties of learning about and making effective use of GIS, the software is in general far easier to use today, and far more productive, than it was in 1984.

The Network

It is amazing to consider that the World Wide Web (WWW), the much-hyped engine of electronic commerce, darling of investors, haven of conspiracy theorists and pornographers, and harbinger of massive changes in our educational system, is only five years old, and was invented only ten years ago by a physicist looking for better ways to share information with colleagues. Today, it is common to talk about the WWW as a vast information resource, and to suggest that we are drowning in a flood of information partly as a consequence of its power and popularity. Certainly, the WWW is having and will continue to have a powerful impact on the development of GIS that will continue over the next 15 years.

Consider for a moment a comparison between the WWW and the average research library, such as exists on my own campus. A typical research library has on the order of 10^6 books in its collection, each containing perhaps 10^5 words, which could be encoded in perhaps 10^6 bytes. Thus, the text in the library might amount to on the order of 10^{12} bytes, or a terabyte. If we assume that the average person reads at 5 words per second, it would take 3,000 years of reading, or 50 lifetimes, to exhaust the library's information resources. Of course, the volume of information in the library is increasing faster than anyone can read, and these figures ignore everything in the library that is not text. In short, the information in the library drowned us long ago. Perhaps it didn't seem that way, because the library is an ordered space of uniform shelves and rows of well-cataloged books, whereas the WWW is a chaotic space of information that is hard to find and may be unreliable when it is found. The library's information has been edited, proofread, and reviewed—it is quality information—and libraries pay collection specialists to ensure that the information in the library is accurate, meaningful, and useful. Libraries also provide the means to find and retrieve information, in ways that are far more sophisticated than the average WWW search engine.

If libraries are so efficient at providing information, what exactly is the function of the WWW? I suggest that the value of the WWW lies in its power as a source of information *not found in libraries.* Although the library has been an excellent mechanism for disseminating information in the form of books, the WWW is clearly better for information that is:

- Timely, and for which the delays in library dissemination due to the lengthy process of writing, publication, and review (which can often take over a year) are unacceptable.
- Hard to handle in traditional form because of problems with the basic medium (photographs, recordings, or maps), or because of problems of cataloging, and where digitization removes these problems.
- Not of general interest (this might include information of personal or local interest), and therefore of little interest to publishers because the economics of the publishing industry favor production in large numbers.

From this perspective, the WWW is an ideal mechanism for distribution and sharing of geographic data. It solves problems of timeliness because, although much geographic information is static, it is difficult for the normal publishing mechanism to deal with updates, corrections, and immediate need. It solves problems of handling, because a map or image in digital form is in principle no more difficult to handle than a book or manuscript in digital form; both are "bags of bits" to a digital network. The WWW also helps solve the problems of cataloging geographic data, because it can process queries about areas on the Earth's surface, something that is very difficult with a conventional card catalog (3). Finally, many types of geographic data fit the third criterion, because detailed data about a local area are not likely to be of major interest in areas outside the immediate region (4).

In the digital world of the WWW, it is possible for anyone equipped with a simple PC to be a publisher. Moreover, advances in geographic information technology, including GIS and the Global Positioning System (GPS), and massive reductions in cost have made it possible for a large number of people and agencies to begin publishing geographic information, despite the fact that much geographic information is of very limited interest. This is revolutionizing traditional arrangements for production and dissemination, which have emphasized central production at the national level and at public expense. Today, WWW-based projects such as the National Geospatial Data Clearinghouse (http://www.fgdc.gov), the Alexandria Digital Library (http://alexandria.ucsb.edu), and Microsoft's Terraserver (http://www.terraserver.com) offer substantial alternatives to the traditional role of the specialized map library, providing information that is more timely and easier to handle, and possibly of very limited interest.

The impact of the WWW on GIS is much greater than its role as a mechanism for dissemination, however. The WWW is an instance of client-server technology, in which operations are divided between a local client provided by the user and a server provided by the host site. Much of the WWW's genius lies in how operations are divided, and in principle it is possible to make use of the WWW from a client that is extremely simple, perhaps nothing more than a $100 add-on to a home television. In such situations, all of the serious computing is done by the server.

Take the example of geocoding, an important function in GIS applications to health problems. Suppose I have a list of 1,000 addresses of patients, and I want to convert them to coordinates in order to map them, or to analyze them in relation to other data. I have two options in today's computing world. First, I could purchase and install a GIS on my desktop, purchase or obtain the necessary data files, enter the addresses, obtain the coordinates, and construct the map. Alternatively, however, I could send the addresses to a WWW site that offers geocoding services, receive the results, and perhaps send them to another site that offers mapping services. The example illustrates the increasingly important distinction between GISystems, as they have been understood for the past two decades, and GIServices, an important and growing area of the WWW. Today, simple GIServices are typically free, financed by advertising revenues. The Mapquest site (http://www.mapquest.com) is an excellent example. Others are more complex, accurate, and timely, and these are services the user should expect to pay for, by providing a credit card or some other straightforward method of electronic payment. Günther and Müller (5) provide an interesting overview of this rapidly developing area of electronic commerce.

How far will GIServices develop, and how much of GIS computing will be transferred to WWW servers? Five years ago, the question was meaningless, but today it is critically important for the future of GIS. What does the concept of GIServices mean for public health? Should agencies be providing GIServices for their clients, as many agencies in other areas already seem to be doing? And how can a public health agency add value to its information by providing services based on it, rather than providing the information itself in raw form?

Software

One of the effects of Moore's Law, and the use of new computing power to build friendlier user interfaces, has been a vast increase in the population of computer users. In 15 years, we have moved from an era in which computing was the preserve of a small elite to one in which virtually everyone, from children to senior citizens, expects computing to be accessible and able to do something useful for them in their daily lives without a great expenditure of effort in training. The personal computer has empowered everyone to compute, write, calculate, and make maps. Everyone today can install a GIS or go to a WWW site and make maps, and no longer is mapping the preserve of a few trained cartographers. In that sense, GIS is truly destroying cartography, though in other senses it is breathing new life into an old and respected discipline.

Computing has put enormous power in the hands of individuals, and produced massive changes in human behavior. So an all-important question for GIS and public health is: How can we harness this enormous empowerment *to improve public health*? As specialists, we can have much greater impact if we focus on empowering others, rather than on the ways in which GIS helps us do our own jobs.

Trends in GIS

All of the issues discussed in the previous section derived from outside the world of GIS, and yet will affect how GIS develops in the next few years. This section addresses issues and trends that are more specific to GIS. Although the MSC report (2) discusses a large number of these, I have selected four that have significance for GIS and, in particular, its applications in public health.

New Data

Over the next few years, a number of new data sources will be coming online with potential for GIS and public health, and it is important that we take advantage of these opportunities as effectively as possible. First, a new generation of satellites will be producing imagery with a resolution of 1 meter. Remote sensing has already proven useful for detecting conditions of importance to public health, such as breeding areas for disease vectors, but this improvement in spatial resolution will create a host of new opportunities for mapping and monitoring conditions that can only be detected at this level of detail. Many indicators of housing quality become visible, for example, as do other socioeconomic variables such as new housing and other indicators of population.

The GPS is also continuing to have impact on data availability. The development of kinematic GPS and its use in vehicles has already reduced the cost of mapping streets by an order of magnitude. Other types of data may make it possible to develop better indicators of lifetime exposure to environmental risk, and new types of census data may make it possible to maintain much more current perspectives on demographic and socioeconomic conditions than is possible with the current decennial system.

New Software

Within the GIS software industry there is a strong interest in achieving interoperability through the adoption of open standards that allow systems developed by different companies to work together without operator retraining or data reformatting. The Open GIS Consortium (http://www.opengis.org) has spearheaded much of this interest, and has developed a number of critically important specifications. It is already possible to operate GIS within Microsoft's Excel, and to open statistical analysis packages without leaving ArcView (ESRI, Redlands, CA), and the indications are that this trend to interoperability will accelerate in the next few years.

One of the benefits of such open specifications is that GIS becomes easier to learn and use, because open specifications require that every vendor adopt the same terminology, or make it possible for someone using another vendor's terminology to use a system without retraining. Open systems inevitably lead to more focus on principles, and less on the details and idiosyncrasies of specific systems. Perhaps surprisingly, in the future, there should be less to learn about GIS.

New GIS-Aware Generations

Thus far, much of the leadership in GIS education has come from the four-year universities, and the majority of courses have been offered at the upper-division or graduate levels. Increasingly, however, the education community has begun to address the needs of other sectors, including non-traditional students who are unable to enter full-time college programs. Distance learning is now well established in GIS through such programs as UNIGIS International (http://www.unigis.org) and ESRI's Virtual Campus (http://www.esri.com). Many community colleges now offer GIS programs, and there are several sources of instructional materials on GIS for this sector (see, for example, http://www.ncgia.org/education/ed.html). GIS is being offered in high schools, and there is interest in exploring its use in elementary schools: imagine, for example, being introduced to the concept of measurement not by using a thermometer to measure temperature, but by using a GPS receiver to measure latitude and longitude.

Over the next decade, the GIS and public health community can expect a much higher level of GIS awareness in its new recruits, and much greater accessibility to GIS functions and expertise.

New Hardware

We have grown used to the idea of computing on a desk, either in the office or at home; for years, computers have been tied to sources of power, and now to Internet connections that only exist where there are phone lines. But it is now possible to compute with full desktop functionality using a portable laptop operating on batteries, and downsizing in the industry is now making it possible to compute on palmtop computers and in vehicles. Moreover, the growth of the wireless communication industry has made it possible to connect from anywhere, and we are rapidly entering the world of mobile, ubiquitous computing, in which location is no longer constrained. We can download data into the field, and upload field-collected data to the Internet or to the office.

Field computing is likely to make a major impact on public health, since it will permit a range of new and exciting opportunities:

- The ability to analyze information as it is collected, rather than later, when the field worker returns to the office.
- The ability to download patient records and other background information to onsite interviews.
- The ability to manage emergencies in the field, on site, and yet have full access to the background information that is needed for effective decision-making.

Strategies for GIS and Health

Given these trends, what can we as a community do to advance the use of GIS in solving public health problems, and to improve public health using these remarkable technologies? I would like to make six suggestions.

Education

First, I suggest that we have to do more to prepare the next generation of public health professionals for GIS, and to raise the awareness of the current generation, much of which was educated before the advent of GIS. We can do this by promoting instruction in GIS in public health schools and by developing partnerships in education with other disciplines that already have well-developed GIS programs: geography, computer science, and geomatics, to name a few. We can promote instruction in GIS at all levels of the education system, with workshops for teachers on GIS applications in public health. We can provide similar opportunities for professionals, through informal education, part-time education, distance learning, the WWW, and workshops at conferences like this.

Research

Second, I suggest there are specific research issues that, if addressed, can improve the use of GIS in public health. Public health applications often require a local focus and the use of local data at the individual level rather than the highly aggregated data that have characterized many previous applications of GIS. They often require an approach that is exploratory, visual, and intuitive. (Anselin [6] has made an excellent review of exploratory spatial data analysis in GIS.) Much health information is uncertain, incomplete, or inaccurate (see, for example, the chapters on data quality in the recent compendium by Longley et al. [7]), and analysis often must be conducted at multiple levels of aggregation.

Data

The WWW is a wonderful resource, but it is not by itself the solution to the need for data in GIS public health applications. Effective searching over the WWW requires the creation of catalog information—metadata, in the language of GIS data access—to enable users to find data more easily and, once they find it, to assess the fitness of data for a given application (3). The National Geospatial Data Clearinghouse is a very effective mechanism for finding data at the national level, but similar efforts need to be promoted at state and local levels, and with the specific needs of public health in mind.

Hardware

Advances in GIS have always relied on advances in hardware generally, and tools like the plotter and the tablet digitizer have given the field very effective boosts in the past. Wireless communication and portable devices seem set to provide comparable opportunities in the future, and there will be other advances we have not even thought about. We need to watch for new opportunities in information technology, and think about how they can improve applications of GIS to public health problems.

Software

GIS is a huge application of information technology, responsible for perhaps $10 billion annually in the United States. Public health is one of many applications of GIS; for GIS vendors, it represents a niche market that may grow into a very significant proportion of the overall market, but as yet is relatively small. One GIS size may not fit all applications, and it is already clear that public health applications present particular needs. We need to promote the development of specialized GIS for public health, and to encourage small software developers who may be able to flourish in this niche to provide add-ons to general-purpose software and systems.

Communication

I noted at the beginning that the potential community of people interested in GIS applications in public health was much larger than this conference, perhaps by two orders of magnitude. We need to find ways to reach that wider audience, and a single national conference cannot possibly do that, given the restrictions on travel that most public health workers face. Too many public health workers see GIS as one of many interesting areas competing for their attention, and conferences like this are at best forums for discussion among national-level agencies and specialists (only 15% of the participants are from state agencies and only 15% are from local agencies, though these figures bear no relationship to the real sizes of these sectors). The solution, it seems to me, is to promote regional and local conferences and workshops in addition to this national forum, perhaps through existing state- and regional-level GIS conferences and organizations, and local chapters of national organizations such as the Urban and Regional Information Systems Association. We should encourage the development of specific public health tracks at these conferences, and also make use of other mechanisms like Chuck Croner's excellent GIS newsletter. (To subscribe to or receive a copy of *Public Health GIS News and Information*, a free bimonthly e-mail report, contact Dr. Charles Croner at cmc2@cdc.gov, or call 301-436-7904, ext. 146.)

The Whole Body Metaphor

I would like to conclude with a point that is somewhat abstract, but nevertheless seems important and an appropriate point with which to end. Metaphors seem particularly useful in trying to address the future, because they allow us to reason and obtain insight through parallels with other fields and concepts. An interesting concept in cognitive psychology is the idea that we learn about the world in childhood by extending concepts from our bodies to our surroundings. Linguists might offer evidence of this in language: "the head of the lake," "the finger lakes," or "the heart of Dixie," while Lakoff (8) has suggested that it is one basis for reasoning about space (and see Mark and Frank [9]).

I suggest that the relationship between GIS and the world is somewhat like the relationship between medical imaging and the human body; a state-of-the-environment report is rather like an individual medical checkup, a report on the state of the geographic body if you like. It is important to the individual if a part of the human body is not working, and it is arguably equally important to society if a part of the world is in bad health. Especially important, if media attention is anything to go by, are problems that are sharply focused in space and time, such as Legionnaire's disease or the Ebola virus.

So my suggestion is that we promote GIS as a tool for exploring the state of the nation's health, with associated diagnostics, policies, and interventions, just as we promote medical imaging as a tool for exploring individual health. The value of a metaphor like this lies, of course, in the thoughts that it provokes: do we need, for example, to develop a profession called "health spatial analyst" that is modeled on the profession of radiology? And what can we learn from the profession of radiology that can help us in imaging the geography of human health?

Acknowledgments

The National Center for Geographic Information and Analysis is supported by the National Science Foundation. Additional support for NCGIA's role in this conference was provided under the Public Health Conference Support Grant Program of the US Department of Health and Human Services.

References

1. Mapping Science Committee. 1993. *Toward a coordinated spatial data infrastructure for the nation.* Washington, DC: National Academy Press.
2. Mapping Science Committee. 1997. *The future of spatial data and society*. Washington, DC: National Academy Press. http://www4.nas.edu/cger/besr.nsf.
3. Goodchild MF. 1998. The geolibrary. In: *Innovations in GIS 5*. Ed. S Carver. London: Taylor and Francis. 59–68.
4. Goodchild MF. 1997. Towards a geography of geographic information in a digital world. *Computers, Environment and Urban Systems* 21(6):377–91.
5. Günther O, Müller R. 1999. From GISystems to GIServices: Spatial computing on the Internet marketplace. In: *Interoperating geographic information systems.* Ed. MF Goodchild, MJ Egenhofer, R Fegeas, CA Kottman. Norwell, MA: Kluwer Academic Publishers. 427–42.
6. Anselin L. 1999. Interactive techniques and exploratory spatial data analysis. In: *Geographical information systems: Principles, techniques, applications and management*. Ed. PA Longley, MF Goodchild, DJ Maguire, DW Rhind. New York: Wiley. 253–66.
7. Longley PA, Goodchild MF, Maguire DJ, Rhind DW. 1999. *Geographical information systems: Principles, techniques, applications and management*. New York: Wiley.
8. Lakoff G. 1987. *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.
9. Mark DM, Frank AU, eds. 1991. *Cognitive and linguistic aspects of geographic space.* Dordrecht, The Netherlands: Kluwer.